

UNIVERSIDAD DE BURGOS

ESCUELA DE DOCTORADO

TESIS DOCTORALES

TÍTULO: MINERÍA DE DATOS Y COMBINACIÓN DE REGRESORES

AUTORA: AGUILAR PARDO, CARLOS
PROGRAMA DE DOCTORADO: INGENIERÍA CIVIL E INDUSTRIAL

FECHA LECTURA: 25/01/2016

HORA: 11:00

CENTRO LECTURA: ESCUELA POLITÉCNICA SUPERIOR. SALA DE JUNTAS 2

DIRECTOR/ES: CÉSAR IGNACIO GARCÍA OSORIO – JUAN JOSÉ RODRÍGUEZ DÍEZ

TRIBUNAL: CARLOS JAVIER ALONSO GONZÁLEZ

JULIÁN LUENGO MARTÍN

GONZALO CERRUELA GARCÍA

JOSÉ FRANCISCO DÍEZ PASTOR

JESÚS MANUEL MAUDES RAEDO

RESUMEN:

La *Inteligencia Artificial* es el área de conocimiento que se dedica a la investigación en la mejora de algoritmos para añadir comportamiento más parecido al humano en los sistemas informáticos. La *Minería de Datos* es una de sus ramas, que está especializada en buscar y extraer información analizando conjuntos de datos. Su objeto de estudio son los sistemas que aprenden por sí mismos. Dentro de los sistemas que aprenden, se llaman *sistemas con aprendizaje supervisado* a aquellos a los que se les proporcionan tanto las entradas como las salidas esperadas de un conjunto de datos de entrenamiento. Cuando la salida esperada es una lista de categorías, el sistema se denomina clasificador, y cuando la salida es numérica, se denomina regresor. La combinación de varios clasificadores o regresores para formar un sistema mejor se suele denominar con el término inglés *ensemble* y suele obtener mejores resultados que el de los métodos independientes que combina.

Esta tesis investiga la mejora que se consigue, al utilizar en tareas de regresión, la combinación de modelos, en vez de la utilización individual de éstos, y analiza que tipo de combinación consigue mayores mejoras. Además, también propone la adaptación a problemas de regresión de algunos algoritmos que estaban diseñados originariamente para clasificación. Más concretamente, se centra en analizar combinaciones homogéneas, es decir, cuando combinan regresores del mismo tipo, e intentar ver qué combinación es la que más mejora a cada tipo de regresor base.

A partir de los resultado de estos estudios se han redactado varias publicaciones: un artículo de revista del primer cuartil, un capítulo de libro, tres comunicaciones en congresos internacionales y otra en un congreso nacional.

Palabras clave: *Minería de Datos*, Regresión, Algoritmos de combinación de modelos, *Rotation Forest*, *Random Oracle*, Proyecciones lineales supervisadas.